DOCUMENT RESUME

ED 149 591 FL 009 145

AUTHOR Charrow, Veda R.
TITLE Let the Rewriter Beware.
PUB DATE 78
NOTE 16p.

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Composition (Literary); *Comprehension; Government Publications; Grammar; Language Research; *Language Styles; *Language Usage; Readability; Semantics; *Syntax; Technical Reports; Technical Writing; Word Frequency; Writing; *Written Language
IDENTIFIERS *Comprehensibility; Flesch Test; Rewriting

ABSTRACT

Translating legal and bureaucratic language into plain, comprehensible English is not amenable to simple rules and procedures. Rewriting comprehensibly requires specialized knowledge about language and an awareness of a number of misconceptions and pitfalls. This paper discusses what not to do in rewriting, based upon rewritten documents presently available to the public and upon several procedures that have been touted as panaceas. Expertise in research design and linguistic analysis is required to indirectly measure or approximate the comprehensibility of discourse. Such indirect measures include multiple choice tests, questionnaires, rating scales, and paraphrase tasks. In contrast, so-called "readability formulas," such as the Flesch Test, which don't measure comprehension or comprehensibility, have become popular. Such formulas only measure sentence length and word length but neglect logical, grammatical, and contextual criteria. Numerous examples are cited to illustrate that rigid rules cannot be made regarding the kinds of grammatical constructions that are most comprehensible. Comprehension problems can arise from a lack of contextual and pragmatic information. Difficulties in rewriting can result from relying on a thesaurus of "word frequency dictionary." Another problem that has arisen in rewriting documents, especially bank loan and insurance forms, is the trend toward personalizing these contracts, which decreases comprehensibility. (Author/SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

© VEDA R. CHARROW 1977

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED149591

Let the Rewriter Beware

Veda R. Charrow, Ph.D.
Center for Applied Linguistics
1611 N. Kent Street
Arlington, Virginia 22209

There has been a growing interest of late in making legal and bureaucratic language comprehensible to the average person. President Carter himself has expressed his concern regarding the incomprehensibility of regulatory language, and has asked for changes in this area. Bills have been introduced and enacted on the federal level, and legislation has been passed on the state level in a number of states to make regulatory language and the language of contracts, insurance policies or bank forms clearer and more comprehensible to the average person. And indeed, bank forms, insurance policies and federal regulations have begun to be rewritten, and professional rewriters and writing consultants of various kinds have appeared in response to the need. Some putative experts have devised their own rules and procedures for rewriting various types of documents; fairly rigid procedures rules and tests which appear to be able to make the rewriter's task an easy, even automatic, one.

Certainly, in this age of mechanized solutions to all sorts of problems, it would seem only natural that there should be automatic and perhaps even automated procedures and rules for translating legal/bureaucratic language into something more understandable to the lay person. Indeed, not only have such procedures been advocated, but many have also been adopted and implemented by some agencies. Unfortunately, translating legal/bureaucratic language into plain, comprehensible English is not amenable to such simple solutions.

FL009145

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Veda Charrow

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

In this paper, I will demonstrate that what appears on the surface to be a rather pedestrian task, is in reality quite complex, and must be undertaken with caution. Rewriting comprehensibly requires specialized knowledge about language and an awareness of a number of misconceptions and pitfalls.

I would have preferred to have written a "how to" paper, explaining in detail how to rewrite comprehensibly. However, such a paper would be impossible, given the complexity of the English language (and of any human language, for that matter) its variability, and the virtually unlimited number of potentially satisfactory writing styles. I will therefore concentrate, in this paper, on what not to do, based upon rewritten documents which are now available to the public, and upon several procedures which have been touted as panaceas.

Before I proceed further, I will briefly deal with the issue of understandability, since it underlies the need to rewrite in "plain English". In order for a legal document to be effective, it must not only be precise and unambiguous, it must also be understood by the persons it is aimed at. If a person has not understood the terms of a contract he has signed, there are certainly logical and moral - if not legal - grounds to question the validity of the contract. But is it possible to measure the understandability (comprehensibility) of a document? And if it is, are there any accurate and simple measures of comprehensibility?

These are not idle questions. There are persons in various fields, such as education and psychology, who have devised tests that purport to measure the comprehensibility of written discourse, but their tests may actually

be measuring something very different. And there are writing consultants who will help to rewrite documents, using rigid and unsophisticated rules and formulas, when it is not altogether clear that they understand what they are doing, or that their product does indeed fulfill the two criteria of understandability and precision.

The truth is, there is no known way to exactly measure the actual comprehensibility of written documents, or actual comprehension on the part of readers/listeners. There are, however, some indirect means of approximating the comprehensibility of discourse - multiple choice tests, questionnaires, rating scales, and the best - paraphrase tasks - but they are not normally used by those who rewrite documents. Such measures require expertise in research design and linguistic analysis; they require time and some expense. A rewriter may have none of these.

To my knowledge, most documents which have been rewritten were neither developed in light of empirical findings of what people do and do not comprehend, nor was their comprehensibility tested by using any of the above-noted methods. Instead, a different sort of measure has been used, and has, for one reason or another, gained popularity as a panacea for rewriters. I am referring to the various so-called "readability formulas", such as the Flesch Test. Indeed, readability formulas have gained such prominence that certain regulations which require that insurance contracts be clear and understandable specify the use of the Flesch Test. What, then, is a "readability formula," and what does it measure?

Readability

To recommend them, readability formulas are very cheap and easy to use, but disregarding for the moment all their other serious limitations, their main drawback is that they do not measure comprehension or comprehensibility. Indeed, they may not even measure that ephemeral and possibly non-existent entity, "readability".

A readability formula is a simple mathematical formula for rating how "readable" a representative passage of written discourse is. Essentially, it involves counting the number of words per sentence, and syllables per word for a given 100-word passage. It works on the premise that the more words per sentence - i.e., the longer the sentence - and the more syllables per word - i.e., the longer the word, - the less "readable" the passage is. The shorter the sentences and words, the more "readable" the passage is. As I will demonstrate, this underlying premise is very dubious. Furthermore there is a large body of linguistic and psycholinguistic data which indicates that comprehension of oral or written discourse has nothing to do with sentence length or word length, but rather depends on some very sophisticated logical, grammatical and contextual criteria. Let me demonstrate what I mean: Since the Flesch test, like other readability formulas, measures only sentence length and word length, two passages with the same average number of words per sentence and the same average number of syllables per word will come out with the same readability score. But sentences can be constructed in a number of different ways which will make sentences of the same length very very different in actual comprehensibility. For example, a long sentence can consist of a number of clauses strung together with "and":

"This morning I got up and brushed my teeth and got dressed and ate breakfast and went to work." (19 words)

But a long sentence might also consist of subordinate clauses embedded one inside the other like a Chinese puzzle:

"The boy who the girl who the man in the white car hit kissed lives next door to me." (19 words)

Both sentences would score about the same on the Flesch test, (the first as somewhat less readable than the second because of the greater number of 2-syllable words) yet the first is so simple, even a beginning reader could understand it, and the second is so hard it takes even a very well educated person a while to figure out what it means.

The following two sentences not only have the same number of words, but they mean pretty much the same thing (although the focus of each is different):

1) The man hit the girl who kissed the boy who eats oatmeal. (12 words)

2) The boy who the girl who the man hit kissed eats oatmeal. (12 words)

And yet again, the second sentence is far less comprehensible than the first. It is the grammatical structure that makes the difference.

Sometimes, as well, a longer sentence is easier to understand and makes more sense than a shorter one. For example:

3) The happening of the accident creates a presumption of defendant's negligence. (10 words)

versus:

4) The very fact that the accident happened allows us to presume that the defendant was negligent. (16 words)

Similarly, word length is not always indicative of comprehensibility. For example, "merry-go-round" has 4 syllables, but even 3-year-olds understand it.

Finally, I could easily build a short sentence, of one-syllable words, which would rate as exceedingly "readable" but which would be totally incomprehensible.

For example: In dog sleep roof street.

"Of course", the reader may say, "that's not a sentence". Precisely. The Flesch test cannot differentiate between sentences and non-sentences; between sense and nonsense; between the comprehensible and the incomprehensible. If a contract were made up of such sentences, the Flesch test would rate it very readable.

Because comprehension is such a large part of one's reading skills, not only do readability formulas such as the Flesch test not measure comprehensibility, they may not even measure "readability." "Readability formulas" should be used very cautiously, if at all - and then only as very gross indicators of overly lengthy sentences. They should not be considered a valid measure of the comprehensibility of discourse, and they should not be used as a criterion for accepting or rejecting rewritten documents.

Grammar

As I mentioned above, comprehensibility does not depend to any great extent upon sentence length. What, then, does it depend upon? It has been suggested that certain aspects of the grammatical construction of a sentence impede comprehension, and that therefore, by altering or eliminating certain grammatical constructions, comprehensibility can be improved. While it is true that certain types of grammatical constructions may affect comprehensibility, these effects may vary from situation to situation. Thus, rigid rules cannot be made regarding what kinds of grammatical constructions

to eliminate. The two types of grammatical constructions most often considered by rewriters to be the culprits are passives (eg., "Fred was sued by Alice" (regular passive); "Fred was sued" (truncated passive)), and subordinate clauses of all types (eg., "The lawyer who won the case collected a large fee"; "He collected a fee which others thought was excessive" (2 embeddings); "I said it was not" (indirect speech), etc.)

Let us see what happens when we eliminate all passives and subordinate clauses in an attempt to improve comprehensibility. Read the following paragraph:

> Fred X is an exceptional person, whose way of life is an example to us all. He is well-liked by most of the people he meets, loved by many, and disliked by very few. He has a lovely wife who adores him, and whom he adores, and his children think he is the most wonderful person who ever lived. Fred is generous to those who are less fortunate than he, without ever being condescending, and at the same time, he is respectful to his superiors without ever being fawning. Before I become too maudlin, let me present this award to Fred X, who we all know to be a man of sterling character and a heart of gold.

Now, read the rewritten version:

> Fred X is an exceptional person. His way of life is an example to us all. People meet him. Most of them like him well. Many love him. Very few dislike him. He has a lovely wife. She adores him. He adores her. His children think something. He is a wonderful person. Wonderful people have lived. He is the

most wonderful person. Fred is generous to people. Those people are less fortunate than Fred. He is never condescending to those people. At the same time, he is respectful to his superiors. He is never fawning. I will become too maudlin. Afterwards, let me present this award to Fred X. We all know something: Fred X is a man. That man has a sterling character. He has a heart of gold.

Obviously, this is an extreme example. But, as you will notice, not only is the "simplified" version childish, and boring, it is actually more difficult to understand than the original. Changing the passives in the second sentence to actives changes the whole focus of the sentence, which is on Fred, not on other people. Removing all the relative clauses also removes logical connectors which are built into complex sentences. The resulting passage makes very little sense.

The point is that although cutting down on the number of passives and subordinate clauses in a passage may be a useful rule-of-thumb, it cannot be done arbitrarily or without regard to the meaning of the sentence. Furthermore, the most common way of eliminating subordinate clauses is by replacing them with prepositional phrases, participial phrases or nominalizations (nouns which are created from verbs - eg. "the implementation"; "the inference"; "the doing"), and this can create its own problems.

eg. Subordinate Clause: When you are incorporating the material into a chapter make sure you know where your typewriter is.

Participial Phrase: When incorporating the material into a chapter, make sure....

Prepositional Phrase: In the procedure of incorporating the material into a chapter, make sure you know ...

Nominalization: The incorporation of the material into a chapter necessitates knowing (participial nominalization) where your typewriter is.

Although the use of nominalizations and various types of phrases often shorten a sentence and cause it to look less complicated than a parallel subordinate clause (and yield "better" readability scores), psycholinguistic studies have shown that these constructions are often more difficult for the average person to process than a subordinate clause. Note that nominalizations tend to be difficult to process, but not merely because they are polysyllabic words; many, such as "the doing (of)", "the asking (of)" are very short. Rather, their difficulty lies in the fact that they have replaced an entire subordinate clause. Nor is it easy to make generalizations as to when such constructions are more difficult; it depends upon the intended meaning of the discourse. In general, however, linguists have found that "reducing" the grammatical complexity in the surface form of a sentence often increases the complexity of the deeper grammatical structure. The meaning of the sentence becomes less clear, and the mind must work harder to decode it.

There are other types of grammatical constructions that tend to be difficult for most people to process, such as conditionals ("if...then"), negative conditionals ("except"; "if not"), multiple negatives (e.g., "He was not unreasonable in denying admittance...."), and others. But although one can postulate certain rules-of-thumb as to their use - such as, in general, avoiding multiple negative - the meaning and the intent of the discourse must always be taken into account first.

Semantics and Context

In addition to grammatical complexity, comprehension problems can arise from a lack of contextual and pragmatic information. That is to say, if the reader does not have a framework into which to fit the new information, - either a framework provided by the context (verbal context) or a framework provided by his own prior knowledge or experience (pragmatic context) - he or she will not be able to understand the new information. The best example of this is from an experiment conducted by the psycholinguists Bransford and Johnson (1972). Read the following paragraph and try to paraphrase it:

> "The procedure is actually quite simple. First you arrange things into different groups depending on their makeup. Of course, one pile may be sufficient depending on how much there is to do. If you have to go somewhere else due to lack of facilities that is the next step, otherwise you are pretty well set. It is important not to overdo any particular endeavor. That is, it is better to do too few things at once than too many. In the short run this may not seem important, but complications from doing too many can easily arise. A mistake can be expensive as well. The manipulation of the appropriate mechanisms should be self-explanatory, and we need not dwell on it here. At first the whole procedure will seem complicated. Soon, however, it will become just another facet of life. It is difficult to foresee any end to the necessity for this task in the immediate future, but then one never can tell."

You have probably fared as badly as the subjects in the original experiment. And yet, if one were to provide a context for the passage, to

wit: "Washing the clothes", immediately the passage becomes comprehensible, and remembering it is no great chore. In the experiment, those subjects who were told the topic before hearing the passage remembered more than twice as many ideas as those who had not been told the topic. They also rated the passage as twice as understandable as those who had not been told the topic. Comprehension and clarity often depends upon having enough context to hang the sentences on. Rewritten forms are often as devoid of context as their predecessors.

Word Frequency

When most persons are confronted with the task of rewriting a passage into "simpler" English, the first thing they tend to do is to try to find "simpler" vocabulary. In order to do this, they may resort to a thesaurus and that useful volume known as the "word frequency dictionary". There are pitfalls here, too.

Word frequency is simply the number of times a given word appears in a large sample of selected reading materials, such as fiction and non-fiction texts, newspapers, magazines, etc. A word which appears frequently - for example, 5000 times in a sample of 5,000,000 words, is a commonly or frequently used word, a high frequency word; because of its high frequency, it is expected that more people would be familiar with it. A word which appears only 10 times in that sample is a low frequency word, a relatively uncommon word. However, depending upon the sample of reading materials chosen for calculating word frequency, and depending upon how the word-frequency is calculated, different word-frequency dictionaries may differ somewhat in their frequency ratings for the same word. Furthermore,

all frequency dictionaries appear to share one major defect: a word is rated for its frequency regardless of its use. Thus, the word "part" can be a noun meaning "portion", "role", or the part in one's hair; it can be a verb meaning to divide something, or to separate from another person. Even though one or two of these very distinct meanings may actually have a very low frequency of occurrence, separate frequencies for each meaning are not available. If a person wanted to find the frequency of the verb "part" meaning "to leave another person", he would not be able to. There is no way of finding the frequency of one meaning of an enormous number of multiple-meaning words such as "part", "can", "voice", (noun and verb), "reason", "talk". This problem is further complicated in the case of what I will call the "less-minus" situation. What happens when a rewriter finds out that a word such as "minus", which he has been using in all his bank contracts, is ten times less frequent (less common) than the word "less"; that, in fact, the word "minus" is a very very uncommon word? If the rewriter assumes, as many people do, that low word-frequency is equivalent to low comprehensibility, he will rewrite the form using "less" in place of "minus". It would be a great mistake to do so. For one thing, "less" has a number of uses and meanings other than "minus": eg. "I want to see less of you" (Not "I want to see "minus" of you); "He eats less than he used to"; "I am less happy than I was". "Minus", on the other hand, seems to be very uncommon, but that is because it has such a very specialized use - as a mathematical operator. The reason its frequency is so low is because it only appears in contexts dealing with the actual use of arithmetic, and most newspapers and books (except for math books - which may not have been chosen for the particular frequency ratings) do

not involve subtraction. For anyone who knows elementary arithmetic, (especially from modern math books) however, "minus" is a much clearer, more precise, and actually easier term than "less". "Less" is ambiguous. On a tax form, for example, when it says "less total deductions", a person who is reading quickly or who reads at a grade school level, might become confused as to whether the "less" means "fewer", "less than", or "subtract from", rather than "minus". A knowledge of how words work, and how frequency dictionaries are made up, is necessary in deciding which word to choose.

It is also necessary, in using a frequency dictionary, to be aware of subtle differences between the meanings of 2 "synonyms" eg. "specific" vs "detailed". If a word is chosen on the basis of its high frequency, rather than its accuracy of meaning, it will cause more problems than it will solve.

Frequency dictionaries can certainly be useful, but they must be used with care and with an understanding of what they actually measure.

Personalization

One of the more egregious trends in rewriting documents - especially bank loan and insurance forms - is that of personalizing these contracts by referring to the client as "I", "me" and the business as "you" - or vice-versa. Although this adds a homey and comforting touch to the documents, it does not necessarily add to their comprehensibility; in some cases, "personalizing" actually detracts from comprehensibility. Again, it is a matter of knowing how to personalize, what purpose it serves, and how it can confuse. Rewriters who have little understanding of workings of language tend to use "personalization" as a panacea, and create new problems in their wake.

For instance, in most cases, the "I" and "you" are defined only once, if at all, at the top of the page, so by the time the reader gets to the bottom of the page, he may well have lost track of "who" is "who". Another problem arises with the use of the pronoun "we, us, our". In a given contract, it may refer to multiple signers (eg. joint tenants), as in "Should any of us die or be incapacitated, the other signers agree to notify you in writing immediately", and "However, you may require all our signatures if there are conflicts among us". Or, it can refer to the signer(s) and the bank, as in "You or I can end our banking relationship at any time". However, the first sentence above could mean that if the loan officer dies, the other signers agree to notify the bank; the second sentence could mean either the bank or one of the joint tenants can end the banking relationship with the other joint tenant at any time, and the third sentence could mean that the loan officer and signers must provide a signature if there is a conflict either among the signers, or between the loan officer and the signers. From the results of my own research into the comprehension of legal language, I am certain that there are borrowers who would understand the sentences in just that way.

In short, personalization is not a bad idea, but it does not necessarily improve a document's comprehensibility, and if used as a cure-all without an understanding of how pronouns work, it can create additional ambiguity.

Conclusion

The growing concern with rewriting legal/bureaucratic documents in clear and understandable English is certainly laudable. But it takes much

more than a set of rigid rules, a thesaurus, a frequency dictionary, and a readability formula to make a document comprehensible. The rewriter should have a good understanding of the way the language works - and, one would hope, some formal training in linguistics - as well as editing experience and a feel for a well-formed sentence. He should also know when not to use mechanical rules and formulas for rewriting, and should fully understand their limitations if he does use them. Ideally, as well, the rewriter should be able to devise and employ valid tests to measure real people's comprehension of the original and rewritten documents. Otherwise, comprehension problems will persist, and consumers will feel they have been cheated with half-measures. As it is, there are no standards for rewriters or writing consultants: persons with dubious qualifications or training can set themselves up as experts. When this happens in other fields, such as law, the offender is guilty of a crime. In the rewriting area, it is the consumer who is punished.